Current Biology 20, R754–R763, September 14, 2010 ©2010 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2010.06.070

Transcriptional Enhancers in Animal Development and Evolution

Review

Mike Levine

Regulatory DNAs serve as templates to bring weakly interacting transcription factors into close proximity so they can work synergistically to switch genes on and off in time and space. Most of these regulatory DNAs are enhancers that can work over long distances — a million base pairs or more in mammals — to control gene expression. Critical enhancers are sometimes even found within the introns of neighboring genes. This review summarizes well-defined examples of enhancers controlling key processes in animal development. Potential mechanisms of transcriptional synergy are discussed with regard to enhancer structure and contemporary ChIP-sequencing assays, whereby just a small fraction of the observed binding sites represent *bona fide* regulatory DNAs. Finally, there is a discussion of how enhancer evolution can produce novelty in animal morphology and of the prospects for reconstructing transitions in animal evolution by introducing derived enhancers in basal ancestors.

Introduction

Shortly after the first glimpse into the molecular organization of eukaryotic genes it became clear that critical regulatory DNAs could be uncoupled from the core promoter, the docking site for RNA Polymerase II (Pol II) [1]. The first eukaryotic enhancers were identified in animal viruses due to their ability to co-opt the transcriptional machinery of host cells upon infection [2–5]. The prototypic enhancer was identified in the SV40 animal virus, an unlikely source for long-range regulatory elements as the SV40 genome is just 5.2 kilobases (kb) in length [1].

The SV40 enhancer contains two 72 base pair (bp) repeats located ~200 bp 5′ of the gene encoding T-antigen, which is essential for viral replication and transcription of late viral genes in infected cells [1]. This 5′-regulatory sequence was shown to work at a distance when attached to a β-globin reporter gene and transfected in cultured monkey kidney cells [1]. Subsequent studies identified several sequence-specific transcription factors that bind to discrete sites within the 72 bp repeats, including the bZIP transcription factor AP1 and the Rel-containing factor NF-κB [6,7]. Both transcription factors are modulated by signaling pathways, receptor tyrosine kinases (RTKs) and Toll, respectively [8,9], and thereby anticipate the importance of enhancers in integrating cell signaling processes — a key insight not appreciated for another decade, e.g., [10]. Indeed, the synergistic activation of the prototypic SV40 enhancer by distinct classes of activators is a common theme of enhancer function and a number of examples will be considered in the course of this review.

University of California, Berkeley, Department of Molecular and Cell Biology, 142 Life Sciences Addition 3200, Berkeley, CA 94720, USA.
E-mail: mlevine@berkeley.edu

Evidence that enhancers might have a more divine purpose than merely augmenting the efficacy of viral infection was obtained by examining genes that exhibit tissue-specific expression. The first cellular enhancers that were identified control the expression of the immunoglobulin (Ig) heavy chain gene in mammalian B lymphocytes [11–13]. Subsequent studies identified a number of sequence-specific transcription factors that bind Ig enhancers, including NF-κB and the basic helix-loop-helix (bHLH) activator E12/E47 [14,15]. Both proteins were subsequently shown to be critical effectors of a variety of processes in animal development and disease, including programmed cell death, inflammation, and lymphocyte differentiation [16,17].

Remote enhancers located tens or even hundreds of kilobases from the target gene are a distinctive property of metazoans that is absent in yeast and rarely seen in plants [18,19]. Such long-range interactions open the door to complex gene control, whereby a given gene can be used in a variety of developmental or physiological processes, as discussed below.

I will first review the general properties of metazoan enhancers, particularly those engaged in developmental processes, and then discuss several well-defined examples.

Principles of Enhancer Function

Activator Synergy

Most developmental enhancers have a number of shared properties, regardless of the overall size of the genomes from which they originate [20–22]. Such enhancers are typically 200 bp to 1 kb in length. They contain multiple binding sites for two or more classes of sequence-specific transcription factors [23]. A recurring theme is the use of at least two different activators to regulate expression, such as NF-κB and AP1 in the case of the SV40 enhancer mentioned above [6,7]. Many enhancers also contain binding sites for sequence-specific repressors, which exclude expression in inappropriate tissues, e.g. [24,25].

Several different modes of transcriptional synergy are known, including cooperative occupancy of linked sites via protein–protein interactions [26,27] and the coordinate recruitment of co-activators such as CBP to the DNA template [28,29]. It is also possible that different classes of activators recruit distinct co-activators, which in turn function synergistically to activate gene expression (Figure 1). For example, activator A might recruit CBP, which mediates acetylation of core histones, while activator B might recruit Swi/Snf, which remodels chromatin by displacing nucleosomes [30,31]. Finally, indirect modes of cooperative binding have been suggested, whereby activator A binds its target site and helps displace the associated nucleosome to facilitate binding of activator B to a neighboring site [32,33] (Figure 1C,D). Most of these mechanisms of activator synergy are non-exclusive, so it is possible, for example, that two activators bind cooperatively to linked sites and coordinately recruit one or more co-activators to the DNA template. Regardless of the exact mechanism, A–B activator synergy depends on close spatial linkage of the binding sites, typically within a half turn of the nucleosome.

The best-characterized enhancer is the enhanceosome controlling the expression of the mammalian β-interferon gene in response to viral infection [34–36]. The enhanceosome is activated by three heterodimeric protein complexes, c-Jun/ATF-2, IRF3/IRF7 and NF-κB, which bind tightly linked recognition sequences that are highly conserved in evolution. A co-crystal of the enhanceosome along with the DNA binding domains of the three activator complexes reveals an extended composite protein surface spanning the length of the ~200 bp enhanceosome [37]. It has been suggested that this surface provides a template for the effective recruitment of co-activators such as CBP (Figure 2).

DNA Looping

Once an appropriate combination of sequence-specific activators and co-activators has bound to an enhancer how does it control gene expression? There is considerable evidence that active enhancers loop to the promoter regions of target genes [38]. These loops can be visualized using chromosome conformation capture (3C) assays and various modifications thereof [39]. These methods are based on identifying the joining of distant DNA segments via PCR.

The exact relationship between the binding of activators and enhancer looping is unclear [40,41]. The *Drosophila melanogaster* Pax2 enhancer contains a discrete region that is important for activation at a distance, but not when the enhancer is placed immediately adjacent to a reporter gene [42]. Cohesins have been implicated as mediators of long-range looping [43,44]. Once the enhancer loops to the promoter it might stimulate transcription by the recruitment of general transcription factors such as the Mediator complex (MED) [45]. In cases of genes containing paused Pol II, activation might be achieved by the recruitment of transcriptional elongation factors such as pTEFb [46].

Enhancer–Promoter Interactions

Whole-genome assays suggest that genes are often contained within larger chromosomal domains that are flanked by insulator DNA [47,48]. Insulators — also known as 'chromosomal barrier elements' or 'boundary elements' — prevent enhancers located within one chromosomal domain from inappropriately activating genes located in neighboring domains [49–55]. In principle, chromosomal inversions or deletions can result in the rearrangement or loss of insulator DNA, and as a result, genes can acquire novel patterns of expression due to activation by enhancers that are normally located within a separate domain.

In some cases, enhancers can bypass a nearby gene in order to activate a more distal transcription unit. This is seen in the *Drosophila* Antennapedia complex, whereby the T1 enhancer is located 3′ of the *fushi tarazu* (*ftz*) transcription unit but bypasses *ftz* in order to activate the more distal *Sex combs reduced* (*Scr*) gene [56–58]. Selective T1–*Scr* interactions depend on a proximal tethering element located immediately 5′ of the *Scr* core promoter (Figure 3A). T1 can activate gene expression from the *ftz* promoter upon insertion of the ~300 bp tethering element. This type of enhancer–promoter specificity is also influenced by sequence elements contained within core promoters. TATA-containing promoters might preferentially interact with certain enhancers, while DPE-containing promoters are activated by different enhancers [59–61].

Tethering elements, and promoter specificity, might enable remote enhancers to ignore nearby genes and activate distal transcription units. These and other mechanisms lead to remarkable examples of uncoupling of enhancers from their target genes. For example, the *sonic hedgehog* (*Shh*) gene is essential for the patterning of the neural tube and limbs [62]. These processes are controlled by separate enhancers, and those regulating expression in the notochord and floor plate are located relatively close to the *Shh* transcription unit. However, the enhancer that regulates *Shh* expression in the developing limb buds is located nearly one megabase away, within the intron of a neighboring

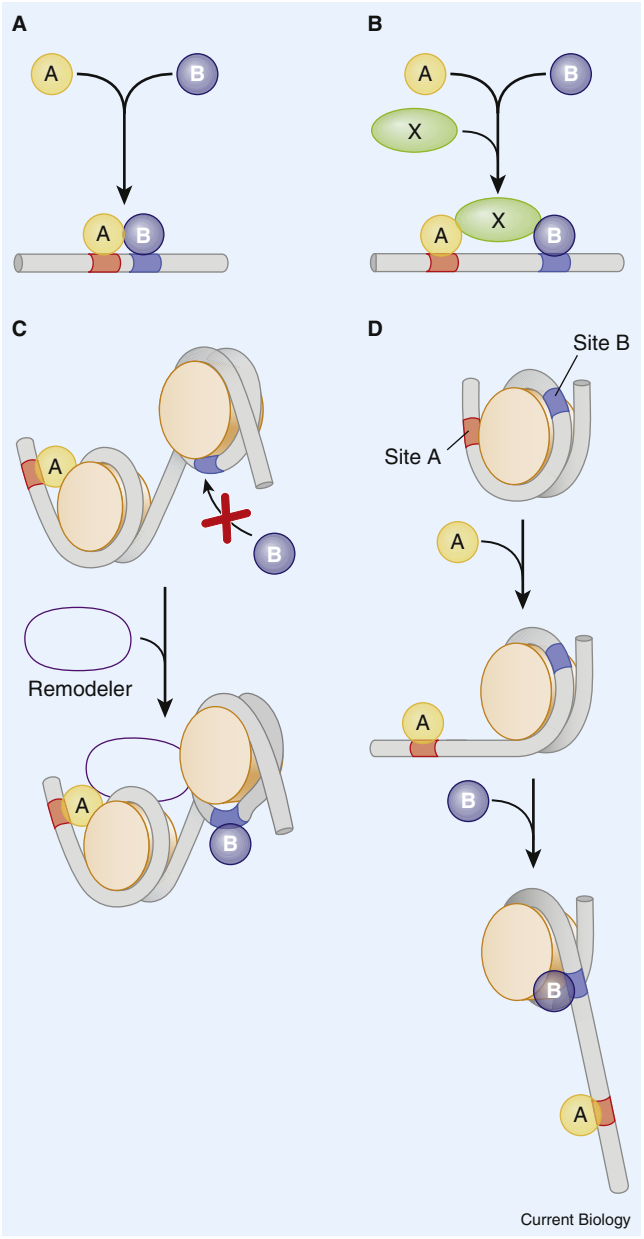

Figure 1. Activator synergy.

Several nonexclusive mechanisms can be envisioned by which two transcriptional activators augment each other's function. (A) Activators A and B cooperatively bind to linked sites. (B) A and B coordinately bind a third protein, X, which stabilizes the binding of A and B. (C) A recruits a histone remodeling protein that facilitates the binding of B. (D) A leads to 'slippage' of the nucleosome and thereby uncovers the B binding site. Reproduced with permission from [125].

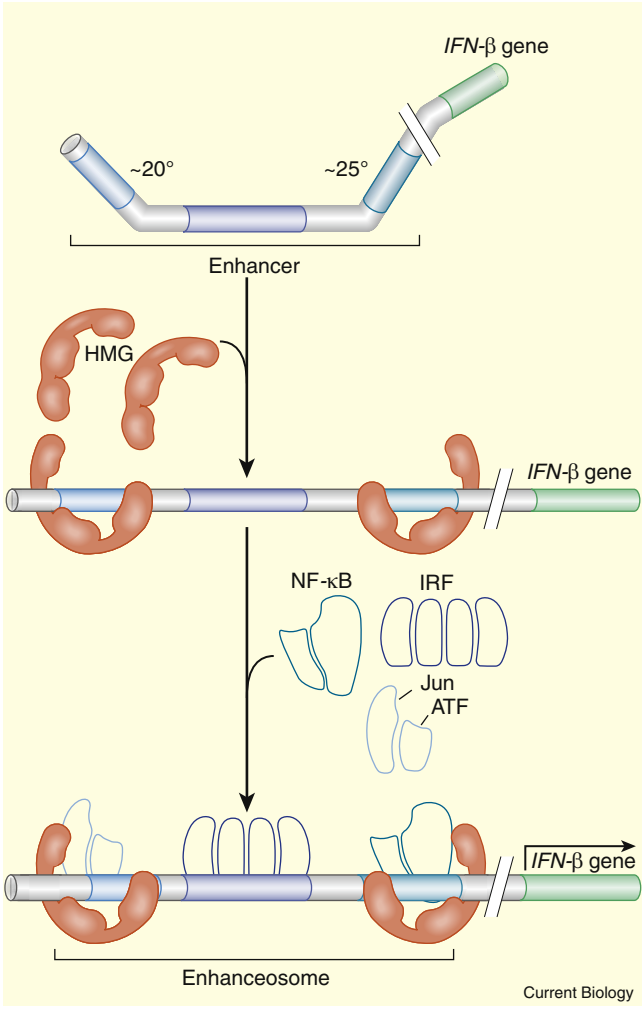

Figure 2. Coordinate recruitment of co-activators at the β-interferon enhanceosome.

HMG bends the enhanceosome and facilitates the binding of NF-κB, IRF, and Jun/ATF to linked sites. The three activator complexes form an extended surface for the recruitment of co-activators such as CBP. Reproduced with permission from [125].

locus (*Lmbr*) [63] (Figure 3B). Once enhancers are decoupled from their target genes, it is thus easy to imagine that they can be re-routed to produce novel patterns of gene expression during animal evolution.

The Question of Grammar

The enhanceosome represents an extreme example of cis-regulatory 'grammar' or 'syntax', as it contains a fixed arrangement of binding sites spanning the entire length of the enhancer [37]. 'Grammar' here refers to the phenomenon that spacing and arrangement of binding sites matter for the activity of the enhancer, just like the order of words in a sentence can affect its meaning. Just about any change in spacing between adjacent sites disrupts enhanceosome function, except the insertion of 10 bp of DNA, one turn of the helix, which maintains the orientation of adjacent heterodimer complexes [34–36]. Developmental enhancers have a more flexible arrangement of binding sites than the enhanceosome. For example, changing the spacing of adjacent Bicoid and Hunchback activator sites does not alter the *even-skipped* (*eve*) stripe 2 expression pattern in *Drosophila*, although there are diminished levels of expression [64]. Nonetheless, it is conceivable that developmental enhancers contain limited arrangements of fixed binding sites.

Just a handful of well-defined developmental enhancers were identified during the 1980s and 1990s. Back in the day, 5′ flanking regions were 'blindly' fused to a reporter gene and the fusion genes were examined in transgenic embryos after incorporation into the germline via microinjection assays [65]. Even so-called 'short-cuts', such as transient expression in injected sea urchin and mouse embryos, required the establishment of demanding microinjection methods [66,67].

The post-genome era provided an opportunity to identify putative enhancers using a host of computational methods, such as phylogenetic foot-printing [68,69], the identification of conserved non-coding sequences among distantly related genomes (e.g., pufferfish and mice). Additional methods include the computational identification of clusters of binding sites [70,71] and whole-genome chromatin immuno-precipitation (ChIP) assays [72,73]. Of course, these methods are not sufficient to identify enhancers, but they can be used to pinpoint the locations of putative enhancers within extended genomic intervals. The combination of computer and experiment has greatly augmented the collection of developmental enhancers, providing a foundation for investigating the question of whether developmental enhancers contain fixed arrangements of binding sites.

The analysis of enhancers that direct gene expression in the presumptive neurogenic ectoderm of the early *Drosophila* embryo suggests a limited grammar in the arrangement of Dorsal and Twist activator binding sites [74]. Dorsal is a Rel-containing transcription factor that is related to mammalian NF-κB, while Twist is a bHLH activator implicated in mesoderm development in a variety of animal embryos [75,76]. Dorsal activates Twist, and low levels of the two proteins work synergistically to activate a number of neurogenic genes (so-called 'type 2 Dorsal target genes'), which restrict EGF and Dpp (TGFβ) signaling in the early embryo [74].

Although the six known type 2 enhancers are unrelated by simple sequence homology, they all contain closely linked Dorsal and Twist binding sites, and the asymmetric Twist site is oriented towards the adjacent Dorsal site. This arrangement fosters cooperative interactions between Dorsal and Twist, but only a subset of the binding sites displays such linkage. Recent computer simulations have been used to argue that linked binding sites might represent evolutionary 'mirages' [77]. However, this example of grammar does not depend on evolutionary constraint, but is seen as a convergent design feature among unrelated enhancers. Moreover, the importance of grammar in the interferon enhanceosome is clear-cut.

Thus, there is no denying that grammar occurs in certain enhancers. The question is whether it is a pervasive feature of developmental enhancers. Later in the review I will describe how changes in the arrangement of binding sites in the *Drosophila* Pax2 enhancer alter gene expression in the *Drosophila* eye [42].

The Proposed Importance of Off-Rates

Sequence-specific DNA binding proteins scan the DNA double helix and then dwell at favored recognition sequences [78,79]. The primary determinant of a 'good' or

high-affinity binding site as opposed to a low-affinity site is the off-rate. DNA binding proteins dwell longer at preferred recognition sequences as compared with sequences that deviate from the optimal consensus sequence.

It is reasonable to suppose that critical recognition sequences must be stably occupied within an enhancer in order to augment transcription. Stable interactions are likely to be required for the recruitment of co-activators, such as CBP, as well as for other aspects of enhancer function such as looping to the core promoter. Stable occupancy depends not only on the intrinsic quality of the binding sites but also on protein–protein interactions, including cooperative binding and interactions with co-activator proteins. It is unlikely that current whole-genome methods, such as ChIP-sequencing, can distinguish between binding sites based on off-rates [80]. Such methods fail to provide kinetic measurements, but represent average states of binding site occupancy. Authentic binding sites responsible for the control of gene expression might depend on regulatory grammar such as protein–protein interactions at linked sites, as discussed earlier.

Function of Individual Enhancers in Animal Development

Below I describe a few specific examples of enhancers mediating localized patterns of gene expression during animal development. There is a particular emphasis on the role of enhancers as DNA templates for integrating complex positional information, including localized signaling molecules.

Xenopus *Organizer*

The dorsal-ventral patterning of the *Xenopus laevis* embryo is controlled by the famous Spemann-Mangold organizer, located at the dorsal lip of the blastopore [81]. The cells comprising the dorsal lip secrete a variety of BMP/TGFβ signaling antagonists, including Noggin and Chordin [82]. The homeobox gene, *Goosecoid* (*Gsc*), is one of the first genes to be activated in the presumptive organizer, where it is thought to participate in the subsequent regulation of patterning genes such as Noggin [83].

The *Xenopus* egg contains a number of localized maternal mRNAs, including Vg1 and VegT, the latter of which encodes a T-box transcription factor that activates Nodal-related genes (XNRs) [84,85]. VegT–XNR interactions create localized activation of pSmad transcription factors in vegetal regions of the early embryo. As pSmads become activated in the vegetal embryonic cells (blastomeres) a process of cortical rotation leads to the localized activation of β-catenin and the Lef/Tcf transcription factor along the presumptive dorsal surface of the early embryo [81,82]. Activated Tcf triggers the expression of the *Siamois* homeobox gene in these dorsal regions [86]. Blastomeres located at the intersection of the vegetal and dorsal regions express both *Siamois* and pSmads (Figure 4). These two sequence-specific transcription factors interact with a 5′ enhancer in the *Gsc* locus to activate *Gsc* expression and thereby help delineate the organizer [87].

Ciona *Heart Primordium*

The beating heart of the adult sea squirt *Ciona intestinalis* arises from a single pair of blastomeres (B7.5) at the 110-cell stage of embryogenesis [88]. A key step in the specification of the heart is the restricted expression of the bHLH regulatory gene *MesP* in the B7.5 blastomeres. *MesP* activates a number of subordinate genes that are required for the directed migration and differentiation of the cardiomyocytes in the developing tadpole [89].

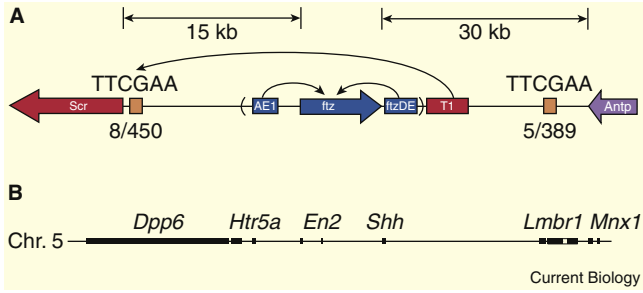

Figure 3. Long-range enhancer–promoter interactions.

(A) The T1 enhancer bypasses the *ftz* locus to activate *Scr* within the Antennapedia gene complex of *Drosophila*. Reproduced with permission from [58]. (B) The primary enhancer controlling *Shh* expression in the developing vertebrate limb bud is located within the intron of the *Lmbr1* locus (unfilled box), which maps nearly one megabase from the *Shh* transcription start site. Reproduced with permission from [63].

MesP is regulated by a 5′ enhancer that contains linked binding sites for two distinct activators, Lhx3 and Tbx6 [90]. Lhx3 is expressed throughout the presumptive endoderm and extends into the B7.5 blastomere, which is located at the boundary between the endoderm and presumptive tail muscles. Tbx6 is expressed in the presumptive tail muscles as well as B7.5. Thus, only the B7.5 blastomeres express both Lhx3 and Tbx6. An outstanding question, posed earlier, is whether a fixed arrangement of Siamois/pSmad and Lhx3/Tbx6 binding sites are essential for the accurate expression of *Gsc* in the *Xenopus* organizer and *MesP* in *Ciona* cardiomyocytes.

Drosophila *Eye*

Pax2 is essential for the specification of cone cells within the adult fly eye. It is regulated by a 362 bp enhancer located ~850 bp upstream of the *Pax2* transcription start site [10,42]. Localized expression within presumptive cone cells depends on Notch and EGF signaling, along with cellular determinants, including Lozenge (Lz), a Runx-containing transcription factor. The *Pax2* enhancer contains a series of 12 binding sites for all three critical regulators: Su(H) (Notch signaling), Ets (EGF signaling) and Lz. Altering the arrangement of binding sites causes the modified enhancer to be active in the R1 and R6 photoreceptor cells rather than cone cells [42]. These experiments clearly illustrate the importance of enhancer structure, or grammar, in the cell-specific regulation of *Pax2* expression. However, there is no evidence for long-range arrangements of binding sites along the length of the enhancer. Rather, there is rapid turnover of binding sites in the *Pax2* enhancers of divergent drosophilids, such as *D. pseudoobscura* [42]. It would appear that grammar may be limited to just a subset of closely linked sites, as seen for the Dorsal/Twist linkage in type 2 neurogenic enhancers.

Caenorhabditis elegans *Nervous System*

The gustatory (taste) neurons of the nematode *C. elegans*, ASER and ASEL, are specified by a zinc finger transcription factor, CHE-1 [91]. A combination of conventional genetics and microarray assays identified a number of putative CHE-1 target genes that are specifically expressed in the taste neurons. The 5′ regulatory regions of a number of these

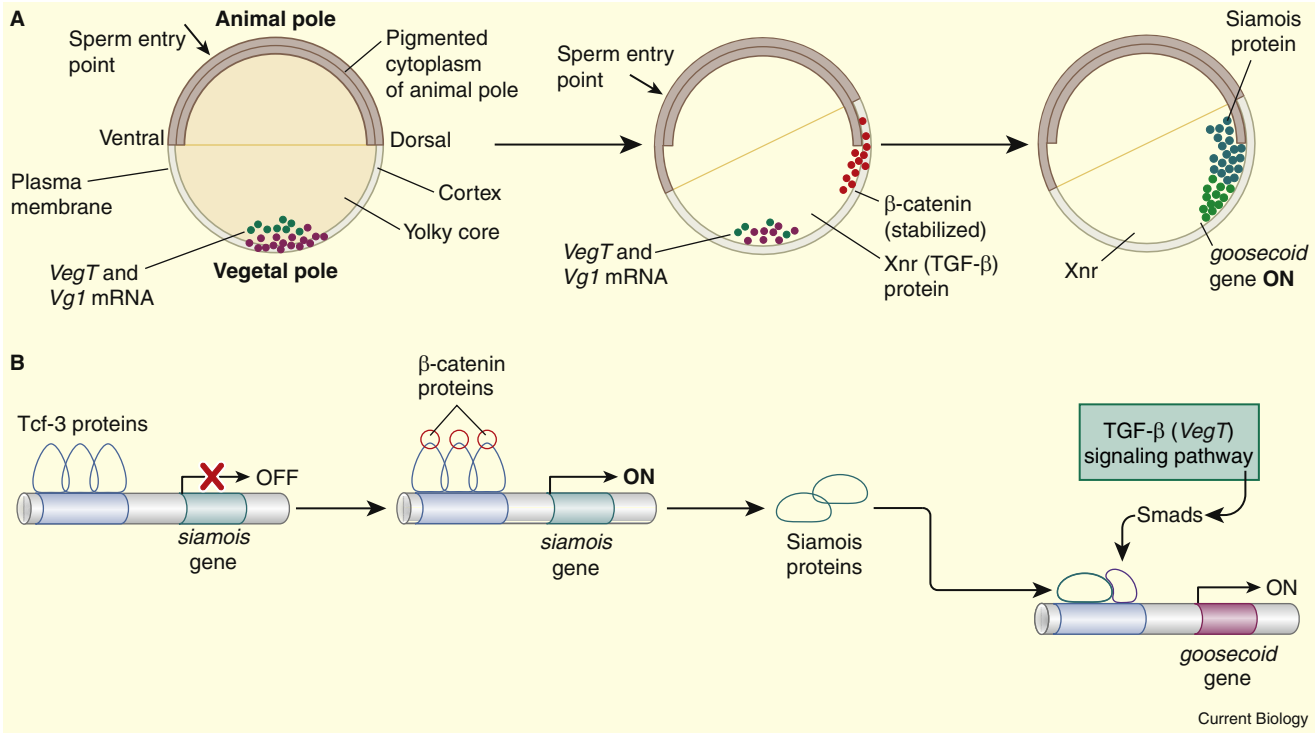

Figure 4. Overlapping activators define the organizer.

(A) *Xenopus* blastula. Xnrs (*Xenopus* Nodal-related signaling molecules) and pSmads are expressed in vegetal regions of the embryo. The homeobox gene Siamois is selectively expressed along the presumptive dorsal surface of the embryo. The two activators, pSmads and Siamois, are co-expressed in the region of the presumptive organizer where they activate the expression of the *Goosecoid* (*Gsc*) gene. (B) Summary of *Gsc* regulation in the organizer. The Smads and Siamois interact with *Gsc* regulatory sequences and activate expression. Reproduced with permission from [125].

genes contain a critical *cis*-regulatory element that is essential for their expression in taste neurons [91]. The CHE-1 target gene *cog-1* contains two copies of this element [92], which might work in a partially redundant fashion to ensure robust expression of the gene in response to environmental fluctuations, as seen for 'shadow enhancers' in *Drosophila* (see below). It is unlikely that the CHE-1 recognition sequence is sufficient for the activation of gene expression in taste neurons. Many target genes exhibit asymmetric expression in the left-right pair of taste neurons, ASEL and ASER, respectively. For example, *cog-1* is specifically expressed in the ASER neuron, and the 5′ regulatory region contains conserved binding motifs that are likely to work in concert with CHE-1 to generate a restricted expression pattern.

Examples of Modular Enhancers in Development

While the preceding examples focused on the regulation of gene expression by individual enhancers, I now consider examples of genes that are regulated by multiple enhancers.

Drosophila *segmentation*

The expression of pair-rule genes in two-segment wide stripes in the *Drosophila* blastoderm embryo provides one of the most dramatic examples of combinatorial gene control in animal development [93,94]. The pair-rule stripes arise after broadly expressed maternal activators (Bicoid and Caudal) and localized gap repressors (Hunchback, Krüppel, Knirps and Giant). The first hint regarding how these broadly distributed regulatory factors produce sharp stripes of gene expression came from the analysis of the 500 bp *eve* stripe 2 enhancer [95,96]. This enhancer contains twelve binding sites; six activator sites and six repressor sites. Bicoid and Hunchback bind the activator sites to augment *eve* expression throughout the anterior half of the embryo (where the two proteins are present at highest levels). The Giant and Krüppel gap proteins bind to the 6 repressor sites to define the anterior and posterior borders of the stripe 2 pattern, respectively [64,97].

eve is regulated by 5 separate enhancers located upstream and downstream of the transcription unit [98]. Each of these enhancers employs the same basic logic described for the stripe 2 enhancer: broadly distributed activators working together with localized gap repressors to define the stripe borders. Three of the gap repressors, Krüppel, Knirps, and Giant, interact with a common co-repressor protein, CtBP, which functions as a short-range repressor [99]. That is, CtBP must bind within 50–100 bp of upstream activators or the core promoter in order to inhibit gene expression. Such short-range repression ensures that the five eve enhancers work independently of one another to produce the complete seven-stripe pattern of gene expression.

Mouse Hindbrain

Hox genes control the anterior-posterior patterning of most or all metazoans [100]. They are usually located within tightly regulated complexes or clusters. Both local and long-range *cis*-regulatory DNA elements control complex patterns of Hox expression during development. The best-characterized vertebrate Hox enhancers are those

controlling the expression of *Hoxb1* in the mouse hindbrain [101,102]. A 3′ enhancer located downstream of the *Hoxb1* transcription unit activates expression in the neurogenic ectoderm, in a broad domain spanning rhombomeres 3 through 5 (R3–R5). A 5′ silencer element containing retinoic acid response elements and Krox20 binding sites represses *Hoxb1* expression in R3 and R5 and restricts expression to the R4 rhombomere.

Hoxb1 expression is subsequently maintained in R4 via a 335 bp 5′ enhancer that mediates auto-regulation. The 5′ auto-regulatory enhancer (R4 ARE) contains a series of linked binding sites for the Hoxb1 protein and a 'ubiquitous' homeobox activator, Pbx. Hox–Pbx protein–protein interactions are thought to result in cooperative occupancy of the linked binding sites, and also foster synergistic recruitment of co-activators [103]. *Hoxb1* regulation is reminiscent of *eve* stripe 2 regulation, in that both systems employ spatially localized repressors for delineating the limits of gene expression. However, the retinoic acid repressor functions over long distances, nearly 2 kb, to silence the Hoxb1 transcription unit in the R3 and R5 rhombomeres.

In addition to local enhancers that regulate individual Hox genes, such as *Hoxb1*, it appears that at least some vertebrate Hox clusters are coordinately regulated by remote sequences that have the properties of the locus control region (LCR) controlling the temporal order of globin gene expression during hematopoiesis [104,105]. The global control region (GCR) of the *Hoxd* complex ensures the sequential expression of the individual *Hoxd* genes along the anterior-posterior axis of developing limbs. There is no evidence that such long-range elements regulate Hox genes in invertebrates such as *Drosophila*. LCR/GCR regulatory DNAs might thus be a distinctive innovation of vertebrates.

Evolution of Gene Expression Patterns

There is growing evidence that changes in gene expression underlie the evolution of novelty and divergence of animal morphology, although there are warring schools of thought regarding the molecular basis for animal diversity [106]. There are those who ascribe most such change to sequential modifications in protein coding sequences, while others insist that changes in *cis*-regulatory DNA are the key agent of animal diversity. There is little doubt that both mechanisms contribute, with DNA duplication as a critical driving force for both types of sequence evolution.

A classical example that highlights the contributions of both protein and *cis*-regulatory evolution in animal physiology is seen at the β-globin locus of vertebrates [107]. The β-globin gene, and its flanking regulatory sequences, underwent a series of duplication events. γ-globin genes are selectively expressed during embryogenesis and the encoded proteins possess a high-affinity for oxygen. This permits the mammalian fetus to compete for the oxygen supply. Upon birth, the embryonic globin genes are shut off and the β-globin genes are activated. The adult globin protein has a lower affinity for oxygen, which is important for gas exchange in deep tissues. In this example, DNA duplication and divergence generated both regulatory DNAs and protein coding sequences with related but distinct activities.

Insect Patterning Genes

Developmental control genes often exhibit distinctive expression patterns in divergent insects. For example, the dorsal-ventral patterning gene *sim* is critical for the specification of the ventral midline of the insect central nervous system [108]. In *Drosophila*, the *sim* expression pattern encompasses just a single row of cells spanning the length of the nerve cord. However, in other insects, particularly the honeybee (*Apis mellifera*), the *sim* expression pattern is broader and encompasses several cells in width, resulting in an expanded midline [109].

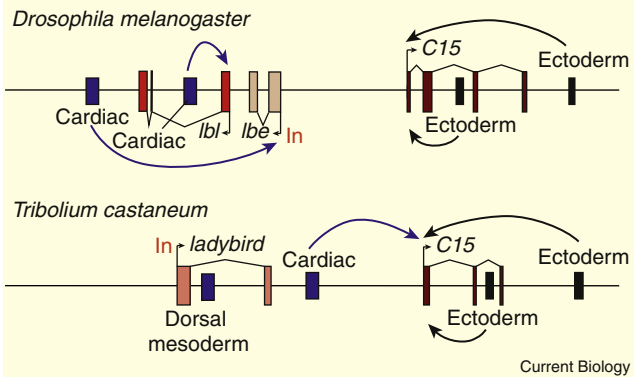

Figure 5. Redirecting a conserved enhancer.

In *Drosophila melanogaster*, 3′ enhancers (blue) activate *ladybird* expression (both the *lbl* and *lbe* genes) in the developing cardiac mesoderm. The *lbe* promoter contains paused Pol II, and has an enhancer blocking activity, preventing the activation of the neighboring C15 gene. In the flour beetle (*Tribolium castaneum*) the single ladybird gene is inverted relative to the orientation of the C15 locus. As a result the 3′ cardiac enhancer is able to activate C15 expression. Reproduced with permission from [118].

In both *Drosophila* and *Apis*, the *sim* expression pattern is controlled by a proximal enhancer located immediately upstream of the core promoter. The *Drosophila sim* enhancer is regulated by a combination of Dorsal, Twist, and Notch signaling. The latter regulatory input restricts *sim* expression to a single line of cells on either side of the presumptive mesoderm (which is the source of the localized Notch signaling protein, Delta). However, the *Apis sim* enhancer contains high-affinity Twist binding sites in place of Notch response elements [109], resulting in an expanded *sim* expression pattern.

'Shadow' enhancers represent another potential source for the evolution of novel patterns of gene expression. Whole-genome ChIP-chip assays suggest that many developmental patterning genes active in the early *Drosophila* embryo contain secondary enhancers, which produce patterns of gene expression that are the same or similar to those generated by more proximal primary enhancers [110,111]. For example, *sog* encodes a secreted BMP inhibitor that is related to *Xenopus* Chordin [112]. The computational identification of clustered Dorsal binding sites identified an enhancer within the first intron of the *sog* transcription unit that recapitulates the endogenous expression pattern in early embryos [70]. ChIP-chip assays identified a cluster of Dorsal, Twist, and Snail binding sites that coincides with this intronic enhancer [110]. These assays also identified a second binding cluster located nearly 25 kb upstream of the *sog* transcription start site, on the other side of a neighboring gene. Despite this remote location, the 5′ binding cluster functions as a *sog* enhancer when tested in transgenic embryos [111]. The shadow enhancer produces a pattern of gene expression that is similar to that seen for the primary, intronic enhancer. Recent studies

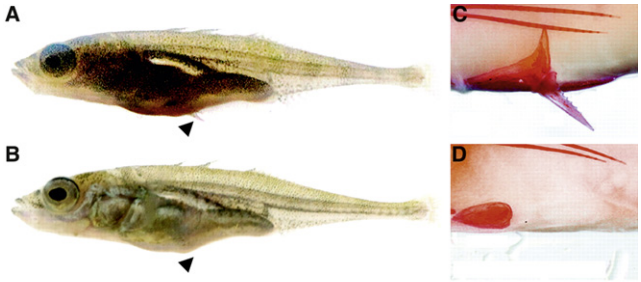

Figure 6. Reconstructing the past.

Ancestral stickleback populations contain pelvic fins (A,C). Certain freshwater populations have reduced fins (B; arrowhead). The fins are restored in these populations upon expression of a transgene containing 2.5 kb of the 5′ flanking region of the Pitx1 locus from a population with pelvic fins. This regulatory sequence was attached to the coding region of Pitx1 derived from a population lacking pelvic fins. Expression of this transgene restores the pelvic fins in populations normally lacking them — compare (C) with transgene to (D) lacking the transgene. Reproduced with permission from [122].

suggest that shadow enhancers might buffer the expression of developmental patterning genes in response to environmental and genetic fluctuations, such as changes in temperature [113–115]. Although fixed in populations by conferring developmental fitness, shadow enhancers might be able to deviate from primary enhancers to produce novel patterns of gene expression at later stages of development.

The preceding examples focus on changes in enhancer sequences that can produce novel patterns of gene expression in evolution. An example of re-directing enhancer–promoter interactions is seen in the Tinman gene complex (Tin-C), which contains a series of homeobox genes that control the patterning of cardial and pericardial cells in the *Drosophila* heart [116]. The *ladybird* gene within the Tin-C is regulated by a 3′ enhancer, which mediates expression in pericardial cells within the developing heart [117]. The *ladybird* promoter region contains paused Pol II and works as an insulator [118]. Consequently, the 3′ enhancer does not activate the C15 gene, which is located upstream of *ladybird* (summarized in Figure 5). However, in the flour beetle *Tribolium castaneum* a chromosomal inversion positions the *ladybird* 3′ enhancer upstream of the C15 gene [118], such that it is now able to activate C15 expression (Figure 5).

Pelvic Fin Reduction in Sticklebacks

While the preceding examples produce only modest changes in morphology, the reduction of pelvic fins in stickleback fish provides a more dramatic example. Sticklebacks are found worldwide in both ocean and fresh water habitats. These different populations have been isolated for over 10,000 years due to the recession of glaciers from the ice age [119]. The ancestral fish contained a pair of prominent pelvic fins, and this is the prevailing phenotype among open ocean populations where vigorous and extended swimming is a mandatory aspect of life. However, a number of isolated fresh water populations have independently loss their pelvic fins. This loss might help sticklebacks escape the grasping jaws of predatory fish.

Crosses between sticklebacks from different populations allowed the genetic mapping of a number of loci responsible for various morphological variations, and a single locus was found to be a major determinant of reduced pelvic fins [119]. This region contains the *Pitx1* homeobox gene, which has been implicated in a variety of developmental processes, including the specification of the mandible, anterior pituitary, and hindlimbs in mammalian embryos. The *Pitx1* regulatory region appears to contain a series of separate enhancers for these various regulatory activities [120].

Pitx1 is selectively expressed in the developing hindlimbs, but not forelimbs. Ectopic expression of Pitx1 in the forelimbs is sufficient to transform forelimbs into hindlimbs [121]. The pelvic fins of sticklebacks are thought to be homologous to the hindlimbs of higher vertebrates. Causal DNA sequence polymorphisms associated with the loss of pelvic fins do not map within the *Pitx1* coding region. Instead, they map within the 5′ regulatory region. Different stickleback populations lacking pelvic fins exhibit deletions of a 5′ ~500 bp enhancer that activates *Pitx1* expression in the developing pelvic fins [122]. This enhancer is located in a 'fragile' region of the genome, near the telomere of linkage group 7. It is therefore possible that the adaptive loss of pelvic fins is facilitated by the fortuitous location of *Pitx1* in the stickleback genome, a location that might foster a high rate of chromosomal deletions.

The modularity of the Pitx1 regulatory DNA is crucial for this loss of pelvic fins without compromising its other activities in development. In mammals, 3′ enhancers are important for different aspects of Pitx1 function, such as the development of the anterior pituitary. These enhancers are unaffected by the chromosomal deletions that remove the 5′ enhancer in sticklebacks.

Reconstructing the Past

The study of pelvic fin reduction in sticklebacks represents a harbinger of the future of 'evo-devo' research, namely, the ability to reconstruct past evolutionary processes in the laboratory. The ancestral pelvic fin enhancer was attached to the *Pitx1* transcription unit and the fusion gene was introduced into a stickleback population containing rudimentary pelvic fins [122]. Remarkably, the transgene was sufficient to restore the pelvic fins in this population, even though they lost their fins over 10,000 years ago (Figure 6). Despite this loss, the entire genetic machinery is competent to respond to the critical expression of *Pitx1* and form fully normal pelvic fins. Thus, it would appear that there has not been significant genetic drift in the regulatory DNAs of the *Pitx1* downstream target genes required for fin morphogenesis.

The loss of stickleback pelvic fins is the premier example of retrograde evolution, the loss of complexity. What is the molecular basis for the acquisition of complexity during evolution? In future years, it might be possible to reconstruct key morphological transitions in animal evolutionary diversity. For example, the gene network underlying the emigration of neural crest cells from the dorsal neural tube of vertebrate embryos is nearly elucidated [123]. It might be possible to promote neural crest formation in an ancestral chordate lacking neural crest, such as amphioxus, by expressing key regulatory genes in the dorsal neural tube (e.g., *FoxD* and *Msx*) [124]. Such genes are present in the amphioxus genome, but they are not normally expressed in the dorsal neural tube as seen in vertebrates. The forced expression of these genes using appropriate *FoxD* and *Msx* enhancers from vertebrates might be sufficient to produce aspects of neural crest formation in amphioxus. Similarly, forced expression of Shh in the fleshy fins of 'transitional' fish might

be sufficient to trigger the development of rudimentary tetrapod limbs.

Acknowledgements

I have always relied on the kindness of strangers and in this regard I thank the anonymous reviewers of the manuscript for correcting many errors and greatly improving the text. I also thank the NIH and NSF for providing research support.

References

1. Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell *27*, 299–308.
2. de Villiers, J., Olson, L., Tyndall, C., and Schaffner, W. (1982). Transcriptional 'enhancers' from SV40 and polyoma virus show a cell type preference. Nucleic Acids Res. *10*, 7965–7976.
3. Schirm, S., Weber, F., Schaffner, W., and Fleckenstein, B. (1985). A transcription enhancer in the Herpesvirus saimiri genome. EMBO J. *10*, 2669–2674.
4. Spandidos, D.A., and Wilkie, N.M. (1983). Host-specificities of papillomavirus, Moloney murine sarcoma virus and simian virus 40 enhancer sequences. EMBO J. *2*, 1193–1199.
5. Hansen, U., and Sharp, P.A. (1983). Sequences controlling in vitro transcription of SV40 promoters. EMBO J. *2*, 2293–2303.
6. Lee, W., Haslinger, A., Karin, M., and Tjian, R. (1987). Activation of transcription by two factors that bind promoter and enhancer sequences of the human metallothionein gene and SV40. Nature *325*, 368–372.
7. Phares, W., and Herr, W. (1991). Functional similarities between human immunodeficiency virus type 1 and simian virus 40 kappa B proto-enhancers. J. Virol. *65*, 2200–2210.
8. Minden, A., and Karin, M. (1997). Regulation and function of the JNK subgroup of MAP kinases. Biochem. Biophys. Acta *1333*, F85–104.
9. Vallabhapurapu, S., and Karin, M. (2009). Regulation and function of NF-kB transcription factors in the immune system. Annu. Rev. Immunol. *27*, 693–733.
10. Flores, G.V., Duan, H., Yan, H., Nagaraj, R., Fu, W., Zou, Y., Noll, M., and Banerjee, U. (2000). Combinatorial signaling in the specification of unique cell fates. Cell *103*, 75–85.
11. Gillies, S.D., Morrison, S.L., Oi, V.T., and Tonegawa, S. (1983). A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell *33*, 717–728.
12. Banerji, J., Olson, L., and Schaffner, W. (1983). A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell *33*, 729–740.
13. Mercola, M., Wang, X.F., Olsen, J., and Calame, K. (1983). Transcriptional enhancer elements in the mouse immunoglobulin heavy chain locus. Science *221*, 663–665.
14. Murre, C., McCaw, P.S., Vaessin, H., Caudy, M., Jan, L.Y., Jan, Y.N., Cabrera, C.V., Buskin, J.N., Hauschka, S.D., Lassar, A.B., *et al.* (1989). Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell *58*, 537–544.
15. Baeuerle, P.A., and Baltimore, D. (1988). Activation of DNA-binding activity in an apparently cytoplasmic precursor of the NF-kappa B transcription factor. Cell *53*, 211–217.
16. Lenardo, M.J., and Baltimore, D. (1989). NF-kappa B: a pleiotropic mediator of inducible and tissue-specific gene control. Cell *58*, 227–239.
17. Murre, C. (2005). Helix-loop-helix proteins and lymphocyte development. Nat. Immunol. *6*, 1079–1086.
18. Levine, M., and Tjian, R. (2003). Transcription regulation and animal diversity. Nature *424*, 147–151.
19. Clark, R.M., Wagler, T.N., Quijada, P., and Doebley, J. (2006). A distant upstream enhancer at the maize domestication gene tb1 has pleiotropic effects on plant and inflorescent architecture. Nat. Genet. *38*, 594–597.
20. Kirchhamer, C.V., Yuh, C.H., and Davidson, E.H. (1996). Modular cis-regulatory organization of developmentally expressed genes: two genes transcribed territorially in the sea urchin embryo, and additional examples. Proc. Natl. Acad. Sci. USA *93*, 9322–9328.
21. Veitia, R.A. (2008). One thousand and one ways of making functionally similar transcriptional enhancers. Bioessays *30*, 1052–1057.
22. Arnosti, D.N., and Kulkarni, M.M. (2005). Transcriptional enhancers: Intelligent enhanceosomes or flexible billboards? J. Cell Biochem. *94*, 890–898.
23. Istrail, S., and Davidson, E.H. (2005). Logic functions of the genomic cis-regulatory code. Proc. Natl. Acad. Sci. USA *102*, 4954–4959.
24. Ip, Y.T., Park, R.E., Kosman, D., Bier, E., and Levine, M. (1991). The dorsal gradient morphogen regulates stripes of rhomboid expression in the presumptive neuroectoderm of the *Drosophila* embryo. Genes Dev. *6*, 1728–1739.
25. Mellerick, D.M., and Nirenberg, M. (1995). Dorsal-ventral patterning genes restrict NK-2 homeobox gene expression to the ventral half of the central nervous system of *Drosophila* embryos. Dev. Biol. *171*, 306–316.
26. Johnson, A.D., Poteete, A.R., Lauer, G., Sauer, R.T., Ackers, G.K., and Ptashne, M. (1981). lambda Repressor and cro–components of an efficient molecular switch. Nature *294*, 217–223.
27. Oehler, S., Alberti, S., and Müller-Hill, B. (2006). Induction of the lac promoter in the absence of DNA loops and the stoichiometry of induction. Nucleic. Acids Res. *34*, 606–612.
28. Dilworth, F.J., and Chambon, P. (2001). Nuclear receptors coordinate the activities of chromatin remodeling complexes and co-activators to facilitate initiation of transcription. Oncogene *20*, 3047–3054.
29. Struhl, K. (2005). Transcriptional activation: mediator can act after preinitiation complex formation. Mol. Cell *17*, 752–754.
30. Malik, S., and Roeder, R.G. (2005). Dynamic regulation of pol II transcription by the mammalian Mediator complex. Trends Biochem. Sci. *30*, 256–263.
31. Ptashne, M., and Gann, A. (2001). Transcription initiation: imposing specificity by localization. Essays Biochem. *37*, 1–15.
32. Schwabish, M.A., and Struhl, K. (2007). The Swi/Snf complex is important for histone eviction during transcriptional activation and RNA polymerase II elongation in vivo. Mol. Cell Biol. *27*, 6987–6995.
33. Shuey, D.J., and Parker, C.S. (1986). Bending of promoter DNA on binding of heat shock transcription factor. Nature *323*, 459–461.
34. Thanos, D., and Maniatis, T. (1995). Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. Cell *83*, 1091–1100.
35. Merika, M., and Thanos, D. (2001). Enhanceosomes. Curr. Opin. Genet. Dev. *11*, 205–208.
36. Agalioti, T., Lomvardas, S., Parekh, B., Yie, J., Maniatis, T., and Thanos, D. (2000). Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter. Cell *103*, 667–678.
37. Panne, D., Maniatis, T., and Harrison, S.C. (2007). An atomic model of the interferon-beta enhanceosome. Cell *129*, 1111–1123.
38. Bulger, M., and Groudine, M. (1999). Looping versus linking: toward a model for long-distance gene activation. Genes Dev. *13*, 2465–2477.
39. van Berkum, N.L., and Dekker, J. (2009). Determining spatial chromatin organization of large genomic regions using 5C technology. Methods Mol. Biol. *567*, 189–213.
40. Zeller, R.W., Griffith, J.D., Moore, J.G., Kirchhamer, C.V., Britten, R.J., and Davidson, E.H. (1995). A multimerizing transcription factor of sea urchin embryos capable of looping DNA. Proc. Natl. Acad. Sci. U S A *92*, 2989–2993.
41. Kostyuchenko, M., Savitskaya, E., Koryagina, E., Melnikova, L., Karakozova, M., and Georgiev, P. (2009). Zeste can facilitate long-range enhancer-promoter communication and insulator bypass in *Drosophila* melanogaster. Chromosoma *118*, 665–674.
42. Swanson, C.I., Evans, N.C., and Barolo, S. (2010). Structural rules and complex regulatory circuitry constrain expression of a Notch- and EGFR-regulated eye enhancer. Dev. Cell *18*, 359–370.
43. Mishiro, T., Ishihara, K., Hino, S., Tsutsumi, S., Aburatani, H., Shirahige, K., Kinoshita, Y., and Nakao, M. (2009). Architectural roles of multiple chromatin insulators at the human apolipoprotein gene cluster. EMBO J. *28*, 1234–1245.
44. Rollins, R.A., Korom, M., Aulner, N., Martens, A., and Dorsett, D. (2004). *Drosophila* nipped-B protein supports sister chromatid cohesion and opposes the stromalin/Scc3 cohesion factor to facilitate long-range activation of the cut gene. Mol. Cell Biol. *24*, 3100–3111.
45. Kornberg, R.D. (2005). Mediator and the mechanism of transcriptional activation. Trends Biochem. Sci. *30*, 235–239.
46. Fuda, N.J., Ardehali, M.B., and Lis, J.T. (2009). Defining mechanisms that regulate RNA polymerase II transcription in vivo. Nature *461*, 186–192.
47. Nègre, N., Brown, C.D., Shah, P.K., Kheradpour, P., Morrison, C.A., Henikoff, J.G., Feng, X., Ahmad, K., Russell, S., White, R.A., et al. (2010). A comprehensive map of insulator elements for the *Drosophila genome*. PLoS Genet. *6*, e1000814.
48. Gaszner, M., and Felsenfeld, G. (2006). Insulators: exploiting transcriptional and epigenetic mechanisms. Nat. Rev. Genet. *7*, 703–713.
49. Kellum, R., and Schedl, P. (1991). A position-effect assay for boundaries of higher order chromosomal domains. Cell *64*, 941–950.
50. Kellum, R., and Schedl, P. (1992). A group of scs elements function as domain boundaries in an enhancer-blocking assay. Mol. Cell Biol. *12*, 2424–2431.
51. Gurudatta, B.V., and Corces, V.G. (2009). Chromatin insulators: lessons from the fly. Brief. Funct. Genomic Proteomic *8*, 276–282.
52. Kuhn, E.J., and Geyer, P.K. (2009). Genomic insulators: connecting properties to mechanism. Curr. Opin. Cell Biol. *15*, 259–265.
53. Chopra, V.S., Cande, J., Hong, J.W., and Levine, M. (2009). Stalled Hox promoters as chromosomal boundaries. Genes Dev. *23*, 1505–1509.
54. Bell, A.C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature *405*, 482–485.
55. Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature *405*, 486–489.

56. Gindhart, J.G., Jr., King, A.N., and Kaufman, T.C. (1995). Characterization of the cis-regulatory region of the *Drosophila* homeotic gene Sex combs reduced. Genetics *139*, 781–795.
57. Calhoun, V.C., Stathopoulos, A., and Levine, M. (2002). Promoter-proximal tethering elements regulate enhancer-promoter specificity in the *Drosophila* Antennapedia complex. Proc. Natl. Acad. Sci. USA *99*, 9243–9247.
58. Calhoun, V.C., and Levine, M. (2003). Long-range enhancer-promoter interactions in the Scr-Antp interval of the *Drosophila* Antennapedia complex. Proc. Natl. Acad. Sci. USA *100*, 9878–9883.
59. Ohtsuki, S., Levine, M., and Cai, H.N. (1998). Different core promoters possess distinct regulatory activities in the *Drosophila* embryo. Genes Dev. *12*, 547–556.
60. Juven-Gershon, T., Hsu, J.Y., and Kadonaga, J.T. (2008). Caudal, a key developmental regulator, is a DPE-specific transcriptional factor. Genes Dev. *22*, 2823–2830.
61. Butler, J.E., and Kadonaga, J.T. (2001). Enhancer-promoter specificity mediated by DPE or TATA core promoter motifs. Genes Dev. *15*, 2515–2519.
62. Vokes, S.A., Ji, H., Wong, W.H., and McMahon, A.P. (2008). A genome-scale analysis of the cis-regulatory circuitry underlying sonic hedgehog-mediated patterning of the mammalian limb. Genes Dev. *22*, 2651–2663.
63. Amano, T., Sagai, T., Tanabe, H., Mizushina, Y., Nakazawa, H., and Shiroishi, T. (2009). Chromosomal dynamics at the Shh locus: limb bud-specific differential regulation of competence and active transcription. Dev. Cell *16*, 47–57.
64. Arnosti, D.N., Barolo, S., Levine, M., and Small, S. (1996). The eve stripe 2 enhancer employs multiple modes of transcriptional synergy. Development *122*, 205–214.
65. Hiromi, Y., and Gehring, W.J. (1987). Regulation and function of the *Drosophila* segmentation gene fushi tarazu. Cell *50*, 963–974.
66. Hough-Evans, B.R., Franks, R.R., Cameron, R.A., Britten, R.J., and Davidson, E.H. (1987). Correct cell-type-specific expression of a fusion gene injected into sea urchin eggs. Dev. Biol. *121*, 576–579.
67. Krumlauf, R., Holland, P.W., McVey, J.H., and Hogan, B.L. (1987). Developmental and spatial patterns of expression of the mouse homeobox gene, Hox 2.1. Development *99*, 603–617.
68. Muller, F., Blader, P., and Straehle, U. (2002). Search for enhancers: teleost models in comparative genomic and transgenic analysis of cis regulatory elements. Bioessays *24*, 564–572.
69. Satija, R., Pachter, L., and Hein, J. (2008). Combining statistical alignment and phylogenetic footprinting to detect regulatory elements. Bioinformatics *24*, 1236–1242.
70. Markstein, M., Markstein, P., Markstein, V., and Levine, M.S. (2002). Genome-wide analysis of clustered Dorsal binding sites identifies putative target genes in the *Drosophila* embryo. Proc. Natl. Acad. Sci. USA *99*, 763–768.
71. Berman, B.P., Nibu, Y., Pfeiffer, B.D., Tomancak, P., Celniker, S.E., Levine, M., Rubin, G.M., and Eisen, M.B. (2002). Exploiting transcription factor binding site clustering to identify cis-regulatory modules involved in pattern formation in the *Drosophila* genome. Proc. Natl. Acad. Sci. USA *99*, 757–762.
72. Lee, T.I., Johnstone, S.E., and Young, R.A. (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. Nat. Protoc. *1*, 729–748.
73. Pepke, S., Wold, B., and Mortazavi, A. (2009). Computation for ChIP-seq and RNA-seq studies. Nat. Methods *6*, S22–S32.
74. Papatsenko, D., and Levine, M. (2007). A rationale for the enhanceosome and other evolutionarily constrained enhancers. Curr. Biol. *17*, R955–R957.
75. Moussian, B., and Roth, S. (2005). Dorsoventral axis formation in the *Drosophila* embryo–shaping and transducing a morphogen gradient. Curr. Biol. *15*, R887–R899.
76. Stathopoulos, A., and Levine, M. (2005). Genomic regulatory networks and animal development. Dev. Cell *9*, 449–462.
77. Lusk, R.W., and Eisen, M.B. (2010). Evolutionary mirages: selection on binding site composition creates the illusion of conserved grammars in *Drosophila* enhancers. PLoS Genet. *6*, e1000829.
78. Berg, O.G., and von Hippel, P.H. (1988). Selection of DNA binding sites by regulatory proteins. Trends Biochem. Sci. *13*, 207–211.
79. Barker, A., Fickert, R., Oehler, S., and Müller-Hill, B. (1998). Operator search by mutant Lac repressors. J. Mol. Biol. *278*, 549–558.
80. Johnson, D.S., Mortazavi, A., Myers, R.M., and Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. Science *316*, 1497–1502.
81. Moon, R.T., and Kimelman, D. (1998). From cortical rotation to organizer gene expression: toward a molecular explanation of axis specification in *Xenopus*. Bioessays *20*, 536–545.
82. Harland, R.M. (1994). Neural induction in *Xenopus*. Curr. Opin. Genet. Dev. *45*, 543–549.
83. Cho, K.W., Blumberg, B., Steinbeisser, H., and De Robertis, E.M. (1991). Molecular nature of Spemann's organizer: the role of the *Xenopus* homeobox gene goosecoid. Cell *67*, 1111–1120.
84. Weeks, D.L., and Melton, D.A. (1987). A maternal mRNA localized to the vegetal hemisphere in *Xenopus* eggs codes for a growth factor related to TGF-beta. Cell *51*, 861–867.
85. Zhang, J., Houston, D.W., King, M.L., Payne, C., Wylie, C., and Heasman, J. (1998). The role of maternal VegT in establishing the primary germ layers in *Xenopus* embryos. Cell *94*, 515–524.
86. Lemaire, P., Garrett, N., and Gurdon, J.B. (1995). Expression cloning of Siamois, a *Xenopus* homeobox gene expressed in dorsal-vegetal cells of blastulae and able to induce a complete secondary axis. Cell *81*, 85–94.
87. Koide, T., Hayata, T., and Cho, K.W. (2005). *Xenopus* as a model system to study transcriptional regulatory networks. Proc. Natl. Acad. Sci. USA *102*, 4943–4948.
88. Davidson, B. (2007). Ciona intestinalis as a model for cardiac development. Semin Cell Dev. Biol. *18*, 16–26.
89. Christiaen, L., Davidson, B., Kawashima, T., Powell, W., Nolla, H., Vranizan, K., and Levine, M. (2008). The transcription/migration interface in heart precursors of Ciona intestinalis. Science *320*, 1349–1352.
90. Christiaen, L., Stolfi, A., Davidson, B., and Levine, M. (2009). Spatio-temporal intersection of Lhx3 and Tbx6 defines the cardiac field through synergistic activation of Mesp. Dev. Biol. *328*, 552–560.
91. Etchberger, J.F., Lorch, A., Sleumer, M.C., Zapf, R., Jones, S.J., Marra, M.A., Holt, R.A., Moerman, D.G., and Hobert, O. (2007). The molecular signature and cis-regulatory architecture of a C. elegans gustatory neuron. Genes Dev. *21*, 1653–1674.
92. O'Meara, M.M., Bigelow, H., Flibotte, S., Etchberger, J.F., Moerman, D.G., and Hobert, O. (2009). Cis-regulatory mutations in the Caenorhabditis elegans homeobox gene locus cog-1 affect neuronal development. Genetics *181*, 1679–1686.
93. Jaeger, J., and Reinitz, J. (2006). On the dynamic nature of positional information. Bioessays *28*, 1102–1111.
94. Levine, M. (2008). A systems view of *Drosophila* segmentation. Genome Biol. *9*, 207.
95. Stanojevic, D., Small, S., and Levine, M. (1991). Regulation of a segmentation stripe by overlapping activators and repressors in the *Drosophila* embryo. Science *254*, 1385–1387.
96. Small, S., Blair, A., and Levine, M. (1992). Regulation of even-skipped stripe 2 in the *Drosophila* embryo. EMBO J. *11*, 4047–4057.
97. Goto, T., Macdonald, P., and Maniatis, T. (1989). Early and late periodic patterns of even skipped expression are controlled by distinct regulatory elements that respond to different spatial cues. Cell *57*, 413–422.
98. Jaynes, J.B., and Fujioka, M. (2004). Drawing lines in the sand: even skipped et al. and parasegment boundaries. Dev. Biol. *269*, 609–622.
99. Zhang, H., and Levine, M. (1999). Groucho and dCtBP mediate separate pathways of transcriptional repression in the *Drosophila* embryo. Proc. Natl. Acad. Sci. USA *96*, 535–540.
100. Lemons, D., and McGinnis, W. (2006). Genomic evolution of Hox gene clusters. Science *313*, 1918–1922.
101. Ferretti, E., Cambronero, F., Tümpel, S., Longobardi, E., Wiedemann, L.M., Blasi, F., and Krumlauf, R. (2005). Hoxb1 enhancer and control of rhombomere 4 expression: complex interplay between PREP1-PBX1-HOXB1 binding sites. Mol. Cell. Biol. *25*, 8541–8552.
102. Ferretti, E., Marshall, H., Pöpperl, H., Maconochie, M., Krumlauf, R., and Blasi, F. (2000). Segmental expression of Hoxb2 in r4 requires two separate sites that integrate cooperative interactions between Prep1, Pbx and Hox proteins. Development *127*, 155–166.
103. Mann, R.S., Lelli, K.M., and Joshi, R. (2009). Hox specificity unique roles for cofactors and collaborators. Curr. Top. Dev. Biol. *88*, 63–101.
104. Mahajan, M.C., Karmakar, S., and Weissman, S.M. (2007). Control of beta globin genes. J. Cell Biochem. *102*, 801–810.
105. Spitz, F., Gonzalez, F., and Duboule, D.A. (2003). A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. Cell *113*, 405–417.
106. Pennisi, E. (2008). Deciphering the genetics of evolution. Science *321*, 760–763.
107. Noordermeer, D., and de Laat, W. (2008). Joining the loops: beta-globin gene regulation. IUBMB Life *60*, 824–833.
108. Crews, S.T. (1998). Control of cell lineage-specific development and transcription by bHLH-PAS proteins. Genes Dev. *12*, 607–620.
109. Zinzen, R.P., Cande, J., Ronshaugen, M., Papatsenko, D., and Levine, M. (2006). Evolution of the ventral midline in insect embryos. Dev. Cell *11*, 895–902.
110. Zeitlinger, J., Zinzen, R.P., Stark, A., Kellis, M., Zhang, H., Young, R.A., and Levine, M. (2007). Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the *Drosophila* embryo. Genes Dev. *21*, 385–390.
111. Hong, J.W., Hendrix, D.A., and Levine, M.S. (2008). Shadow enhancers as a source of evolutionary novelty. Science *321*, 1314.
112. Francois, V., Solloway, M., O'Neill, J.W., Emery, J., and Bier, E. (1994). Dorsal-ventral patterning of the *Drosophila* embryo depends on a putative negative growth factor encoded by the short gastrulation gene. Genes Dev. *8*, 2602–2616.

113. Frankel, N., Davis, G.K., Vargas, D., Wang, S., Payre, F., and Stern, D.L. (2010). Phenotypic robustness conferred by apparently redundant transcriptional enhancers. Nature *466*, 490–493.

114. Perry, M.W., Boettiger, A.N., Bothma, J., and Levine, M. (2010). Shadow enhancers foster robustness of Drosophila gastrulation. Curr. Biol. *20*, 1562–1567.

115. Hobert, O. (2010). Gene regulation: enhancers stepping out of the shadow. Curr. Biol. *20*, R697–R699.

116. Zaffran, S., Xu, X., Lo, P.C., Lee, H.H., and Frasch, M. (2002). Cardiogenesis in the *Drosophila* model: control mechanisms during early induction and diversification of cardiac progenitors. Cold Spring Harb. Symp. Quant. Biol. *67*, 1–12.

117. Liu, J., Qian, L., Han, Z., Wu, X., and Bodmer, R. (2008). Spatial specificity of mesodermal even-skipped expression relies on multiple repressor sites. Dev. Biol. *313*, 876–886.

118. Cande, J.D., Chopra, V.S., and Levine, M. (2009). Evolving enhancer-promoter interactions within the tinman complex of the flour beetle, Tribolium castaneum. Development *136*, 3153–3160.

119. Shapiro, M.D., Marks, M.E., Peichel, C.L., Blackman, B.K., Nereng, K.S., Jónsson, B., Schluter, D., and Kingsley, D.M. (2004). Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. Nature *428*, 717–723.

120. Lanctôt, C., Moreau, A., Chamberland, M., Tremblay, M.L., and Drouin, J. (1999). Hindlimb patterning and mandible development require the Ptx1 gene. Development *126*, 1805–1810.

121. DeLaurier, A., Schweitzer, R., and Logan, M. (2006). Pitx1 determines the morphology of muscle, tendon, and bones of the hindlimb. Dev. Biol. *299*, 22–34.

122. Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G., Jr., Shapiro, M.D., Brady, S.D., Southwick, A.M., Absher, D.M., Grimwood, J., Schmutz, J., et al. (2010). Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. Science *327*, 302–305.

123. Sauka-Spengler, T., and Bronner-Fraser, M. (2008). A gene regulatory network orchestrates neural crest formation. Nat. Rev. Mol. Cell. Biol. *9*, 557–568.

124. Yu, J.K. (2010). The evolutionary origin of the vertebrate neural crest and its developmental gene regulatory network–insights from amphioxus. Zoology (Jena) *113*, 1–9.

125. Watson, J.D., Baker, T.A., Bell, S.P., Gann, A., Levine, M., and Losick, R. (2007). Molecular Biology of the Gene, 6th Edition. Benjamin Cummings.